Precision Agriculture, 6, 431–452, 2005

Site-specific Approaches to Cotton Insect Control. Sampling and Remote Sensing Analysis Techniques

J. L. WILLERS, J. N. JENKINS AND W. L. LADNER jlwillers@ars.usda.gov
USDA ARS, Genetics and Precision Agriculture Research Unit, Mississippi State, MS, USA

P. D. GERARD
Experimental Statistics Unit, Mississippi State, MS, USA

D. L. BOYKIN
USDA ARS, Statistics Unit, Stoneville, MS, USA

K. B. HOOD
Perthshire Farms, Gunnison, MS, USA

P. L. MCKIBBEN
McKibben Ag Services, LLC, Mathiston, MS, USA

S. A. SAMSON
Extension GIS and GeoResources Institute, Mississippi State, MS, USA

M. M. BETHEL
ITD Spectral Visions, Stennis Space Center, MS, USA

Abstract. When insect population density varies within the same cotton field, estimation of abundance is difficult. Multiple population densities of the same species occur because cotton fields (due to edaphic and environmental effects) are apportioned into various habitats that are colonized at different rates. These various habitats differ temporally in their spatial distributions, exhibiting varying patterns of interspersion, shape and size. Therefore, when sampling multiple population densities without considering the influence of habitat structure, the estimated population mean represents a summary of diverse population distributions having different means and variances. This single estimate of mean abundance can lead to pest management decisions that are incorrect because it may over- or under-estimate pest density in different areas of the field. Delineation of habitat classes is essential in order to make local control decisions. Within large commercial cotton fields, it is too laborious for observers on the ground to map habitat boundaries, but remote sensing can efficiently create geo-referenced, stratified maps of cotton field habitats. By employing these maps, a simple random sampling design and larger sample unit sizes, it is possible to estimate pest abundance in each habitat without large numbers of samples. Estimates of pest abundance by habitat, when supplemented with ecological precepts and consultant/producer experience, provide the basis for spatial approaches to pest control. Using small sample sizes, the integrated sampling methodology maps the spatial abundance of a cotton insect pest across several large cotton fields.

Keywords: sampling, probability models, remote sensing, cotton insect management

Introduction

Sampling large cotton fields to determine the status of any insect pest is a difficult and time-consuming task. Traditionally, sampling cotton pests has entailed the visual inspection of plants, sweep or drop cloth samples in a field to obtain data used to estimate overall abundance. Evidence that this strategy is faulty is the fact that insect control ranks among the highest costs in cotton production (Williams, 1998; Dupont *et al.*, 2000; Reed, 2001a) yet yield losses still occur. One of the causes for this situation is a lack of spatially precise information about pest abundance. Today, global positioning systems (GPS), geographic information systems (GIS), remote sensing (RS) and variable-rate (VR) technology are available to create new approaches for pest control. These technologies allow for increases in production efficiency by improving how and where to sample for cotton pests.

Many entomological sampling works emphasize the importance of sample size (*e.g.*, Karandinos, 1976). However, in developing the potential of using information to improve insect control on commercial cotton farms, the emphasis must shift toward sample unit size rather than sample size, especially if remote sensing is available. Byerly *et al.* (1978), Morris (1955) and Southwood (1978) describe several sample unit attributes. They state a sample unit should (1) have an equal chance of selection, (2) be stable, (3) be convertible to unit area, (4) be easily delineated in the field, (5) strike a balance between the required accuracy and cost of collecting the sample, (6) have the proportion of the insect population (over a short time) using the sample unit as a habitat be constant and (7) approximate the average ambit of mobile individuals. Many of these attributes apply to the location captured by the individual pixels of a digital, remotely sensed image of cotton fields.

Different sampling plans have different characteristics (Thompson, 1992). The image-based plan developed here proposes that a sample unit comprised of a single cotton plant is too small. Therefore, a distinctive characteristic is that sample units are small areas of ground formed by assimilating collections of adjacent plants from small lengths of row on one or more adjoining crop rows. The number of consecutive plants comprising a sample unit (and thus its area) differs for different sampling methods. A second characteristic is that effective sample unit sizes are inversely related to the pest population density. Specifically, as pest density increases, the ability of a sampling method to detect at least one insect in a sample unit of any given size will increase. Other characteristics are (1) the various population densities of the same pest species spatially residing within cotton fields correspond to habitat quality and (2) classified, multi-spectral imagery having high spatial resolution (*i.e.*, small pixel sizes) effectively maps the diverse habitats of cotton fields. These habitats define the sampling strata of relevance for a fruit-feeding cotton insect pest, such as the tarnished plant bug (TPB; *Lygus lineolaris* [Palisot de Beauvois] (Heteroptera: Miridae)). The availability of a map showing the locations of these habitats and their association with one another reduces the need to have large sample sizes in any particular habitat class. Another important characteristic is that supplementary information provided by consultant and/or producer knowledge may indicate that not all habitat classes need to have equal sample sizes nor do all parts of a particular class need to be sampled. The last characteristic is that the simple random sampling (SRS) (Thompson, 1992) design (within habitat classes) allows one to estimate pest abundance (λ_f) at a sample location. Also under the SRS design, several samples from a habitat class provide an estimate of its risk to pest attack (p). The SRS design is efficient and practical on commercial farms.

Several articles (Willers *et al.*, 1999; Dupont *et al.*, 2000; Reed, 2001a, b, c) discuss utilizing diverse spatial-based technologies for cotton insect management. However, the acquisition and integration of the information necessary to build the sampling maps (Fleischer *et al.*, 1999) remains undescribed. Within the entomological sampling literature (Southwood, 1978; Pedigo and Buntin, 1994), few works (Pinter *et al.*, 2003) employ spatial technologies with sampling efforts in commercial cotton fields. Our objective is to describe an integrated sampling methodology offering experienced cotton entomologists the flexibility to (1) change the sampling method, (2) determine if additional pest species are present, (3) choose where to sample additional locations and (4) decide when and where treatment actions should be implemented. Several ecological concepts and statistical modeling results are discussed to develop the sampling approach. A field case study, using the tarnished plant bug, illustrates its application.

Methods

Ecological definitions and concepts

Colinvaux (1973) defines a community as a collection of different varieties of organisms living together. By this definition, insect and cotton plant populations within fields aggregate into different communities that vary spatially over time. These field communities are different from the non-crop communities existing outside the field boundary; however, pest and beneficial insects can move between them. Differences in the phenology of the crop establish the habitat quality (Andrewartha and Birch, 1970; Daubenmire, 1974) which influences insect diversity and determines how many communities occur within cotton fields. Variability in habitat quality results from interactions among agronomic practices, soil types, hydrology and weather. Cotton habitat quality can be categorized (by remote sensing methods) to establish spatially distinct populations of cotton plants interspersed within the same field and nearby fields.

For developing site-specific pest management approaches, the various populations of an insect pest are defined by their places of residence and not by gene flow (Krebs, 1978) among individuals. Specifically, a particular habitat class will possess temporally short (weekly), but stable, demographics for the pest population mean, variance and age structure. Other cotton habitat classes will have additional populations of that pest having different demographics. More than a century of ecological thought indicates that animals and birds prefer one habitat to another (with different colonization rates) and cotton insect pests are no different. Therefore, sets of demographic relationships with unique habitats define where various populations of a cotton insect pest reside. Past crop and insect management practices are other effects that can modify pest population demographics.

Probability models and concepts

Negative binomial distribution (NBD) model. Computer model results established the sample unit sizes recommended (see below) for different cotton insect sampling methods. This model was based upon the negative binomial distribution (NBD) (Anscombe, 1949; Pielou, 1977) and investigated the relationship between pest population density and sample unit size under the assumption of random dispersion. To model random dispersion (Willers *et al.*, 1990) of insects within a habitat at a particular density, the dispersion parameter of the NBD was fixed at a large value ($k=50$). The NBD converges to the Poisson (or random) distribution as $k \to \infty$ (Anscombe, 1950; Pielou, 1977) and a value of 50 is sufficiently large for this to be satisfied in practical terms.

The NBD model simulated the probability of finding r insects ($r=0, 1, 2, 3\ldots, 99, 100$) in a sample unit of a specific size ($s \equiv 1$ single plant or 10, 25, 50, 66 or 100 adjacent plants) as:

$$P(r) = \binom{k+r-1}{r} \left[\frac{m}{m+k}\right]^r \left[\frac{k}{m+k}\right]^k \tag{1}$$

where

$k=$ the clustering (or dispersion) parameter (fixed at 50) and
$m=$ the average number of insects per sample unit.

The influence of sample unit size can be modeled if $s^*\lambda_c$ is substituted for m, where λ_c is a rate describing the number of insects per plant. (For example, given a sample unit size of 100 plants and $\lambda_c=0.40$ insects/plant, you would expect to find 40 insects in the sample unit.) In addition, if $\lambda_c>0.40$, the infestation level of the pest among cotton plants is so great, that it is unnecessary to model higher densities. (This paper defines and applies three definitions for an infestation rate. First, for computer runs of different sample unit sizes, λ_c is the infestation rate parameter to model a particular level of pest abundance. Second, the parameter λ_f represents the infestation rate estimated by a sampling method at a sample location. The third parameter, λ (without a subscript), refers to the actual infestation rate of a habitat in a cotton field, which is unknown, and is estimated by sampling).

A SAS® program was written to find $P(r)$ for different sample unit sizes (s) and selected infestation rates ($\lambda_c=0.04$, 0.10 and 0.40). Cumulative density functions (CDF), $F(r)$, were graphed to make comparisons among different combinations of values for λ_c and s with a fixed k.

The model can also simulate insect abundance by creating a grid of sample units of a specified size where the number of insects in each cell is a random variate. The spatial extent of this simulated habitat for a particular model configuration is an abstraction (Willers *et al.*, 1990), while in practice, the geo-spatial extent of a cotton habitat is determined from classified imagery. Agronomically, these simulated sample units (if and only if $s > 1$) represent a specified length of crop row spanning two or more adjoining furrows (Willers and Akins, 2000; Willers *et al.*, 1990) until an area of land is created that contains s plants. In the cotton field, to minimize spatial correlations among plants, it is best to create larger sample unit sizes by assimilating cotton plants from adjoining rows.

Resampling model. Another statistical model illustrated the impact upon estimates of pest abundance if the sampler was unaware of heterogeneous pest densities within a cotton field. This SAS® program employed the inverse transformation method (Pritsker and Pegden, 1979) and the NBD to generate simulated samples. The example assumes the use of a large sample unit size (or $s = 66$ adjacent plants) per sample. The model generated 12 random variates for each case (Table 1) where $\lambda_c = 0.01$ and 0.40, while $k = 50$ remained fixed. One of twelve sample variates from each infestation rate was assigned to a sample unit.

A small, imaginary field was created that was composed of the 24 sample units (specified above) and sampled with a sample size (n) of 4. To model the sampling process, the field units were resampled (Manly, 1997; Willers *et al.*, 2000) using the general formula for the enumeration (Beyer, 1968) of combinations of 24 items taken 4 at a time (${}_{24}C_4$). A frequency histogram (using the chart procedure in SAS®) shows the distribution of mean estimates for insect abundance (in terms of λ_c) for all combinations of four samples.

Image acquisition and processing

Remote sensing acquisition, classification and GIS processing techniques spatially map cotton habitat classes. The imagery used in this project was obtained from an airborne, digital camera system [ITD-Spectral Visions RDACS (Remote Data Acquisition Camera System) Stennis Space Center, MS, USA] that acquired three bands (540 ± 5, 695 ± 5 and 840 ± 5 nm) at 2 m spatial resolution. Image acquisition was similar to that in Willers *et al.* (1999), with the exception that the aircraft flew at 3648 m above ground level (AGL). The imagery was geo-referenced to GPS ground control points, using nearest neighbor resampling, to the Universal Transverse Mercator (UTM) coordinate system in the WGS 1984 datum (ESRI Institute, 1994), acquired on 8 June 1999 and spanned 218 ha of cotton acreage.

Image classification. From the multi-band composite, the normalized difference vegetation index (NDVI) was determined for each pixel in the image frame using ERDAS Imagine® software (Leica Geosytems (ERDAS), Atlanta, GA, USA). The

Table 1. Matrix representation of 24 simulated field sample units infested by a pest insect at two actual rates (λ_c) to prepare data used by the resampling experiment (see text)

Rate	Value											
$\lambda_c = 0.01$	0.00	0.03	0.00	0.00	0.00	0.03	0.06	0.00	0.00	0.03	0.00	0.03
$\lambda_c = 0.40$	0.42	0.45	0.39	0.36	0.21	0.45	0.33	0.36	0.39	0.42	0.33	0.24

The sample unit size for each generated random variate was 66 adjacent plants.

NDVI is a continuous index (−1 to 1) that measures crop vigor. It is a ratio of the difference between the near infrared (nir) and red bands divided by the sum of the same two bands (Rouse *et al.*, 1974).

An unsupervised classification step (Jensen, 1986; Pouncey *et al.*, 1999) using Imagine® first apportioned the NDVI values into 20 classes (using 1 standard deviation and 0.95 as the criteria for the stopping rule [software options]). Supervised classification techniques (Richards and Jia, 1999) next iteratively recoded the colors of the image attribute table to obtain the final classification of landscape categories. Four cotton habitat classes and four other landscape classes were established.

The issue of image calibration or normalization prior to classification analysis is a topic beyond the scope of this presentation (Yuan and Elvidge, 1996; Edirisinghe *et al.*, 2001). However, the image based SRS plan described here does not require radiometrically calibrated or normalized imagery, because a map describing *relative* variability in cotton phenology (as categorical habitat classes) is sufficient. In addition, the NDVI is a ratio of two spectral bands that provides some radiometric adjustment (Jensen, 1986; Thenkabail *et al.*, 2002).

Sampling definitions and concepts

A sample is collected at a particular geographic (x, y) location within a habitat class. The number of samples collected with a particular sampling method during a specific interval of time defines the sample size (n) for a particular habitat class.

Using visual, drop cloth or sweep net samples within a SRS design (Thompson, 1992), the consultant estimates pest abundance at a location within a habitat class using the recommended sample unit size (Table 2) for that method. Calculations for estimating the infestation rate ($\hat{\lambda}_f$) of a sample and the relationships between λ_c and λ_f for different sampling methods are presented. Control methods are necessary at the sample location if the estimated infestation rate equals or exceeds a user specified action threshold (for example, $\lambda_f \geq 0.06$ TPB/sweep). The average infestation rate of a habitat class is the mean of its n samples.

Another quantity estimated under SRS is the risk of a particular habitat class to damage by the cotton pest. This proportion is estimated by $\hat{p}$ and is obtained by dividing the number of samples in a habitat, whose estimated infestation rates are more than or equal to the action threshold, by the sample size (n) of the habitat (Table 2). Estimates of risk lead to a management decision to (1) spray or (2) not spray or (3) collect more samples from a habitat class. Breakpoints or class limits for these decisions are also user specified; for example, if $\hat{p} > 0.7$, then spray the habitat.

Other than the estimate of pest abundance, a sample can have additional attributes. The label (Thompson, 1992) identifies the sample's location as a pair of map coordinates (*e.g.*, latitude and longitude). Auxiliary variables (Thompson, 1992) are additional information about the sample (habitat classification, beneficial abundance, if the estimate exceeds the action threshold (yes or no), the recommended control tactic, the number of days since the last pesticide application, the cotton variety and other attributes).

Table 2. Summary features of several methods applicable to cotton insect sampling

Image-based, stratified, simple random sampling methods					
Sampling method[a]	Size of sample unit (Area)	Sample size (n)/Habitat	s[b]	Rate Estimator/Sample	Proportion (Risk) Estimator/Habitat Class[c]
Visual	2 rows by 2.74 m (ca. 3 m^2)	3–7	50	$\hat{\lambda}_f$ = No. Insects/plant	$\hat{p}$ = No. samples infested/n
Drop cloth	12 Linear row feet (ca. 4 m^2)	3–7	36	$\hat{\lambda}_f$ = No. Insects/row foot	$\hat{p}$ = No. samples infested/n
Sweep net	33 Sweeps (ca. 12.45 m^2)	3–7	66	$\hat{\lambda}_f$ = No. Insects/33 sweeps	$\hat{p}$ = No. samples infested/n

For all methods, the sample unit represents a sample.

[a]In practice, not each plant or individual sweep is examined for insects. Only the total numbers of insects for the sample unit are recorded. However, for the drop cloth, there is useful information if the counts are recorded by row (see Willers and Akins, 2000).

[b]Modal number of adjacent plants per sample unit (Derived from Figure 2, where $\lambda_c = 0.04$ and P(r = 2–4 insects) $\approx$ 0.5). Number of adjacent plants per sample unit establishes the correspondence between λ_c and λ_f. For example, with the sweep net method, (66 plants per 33 sweeps = 2 plants per sweep).

[c]A sample is infested if the number of insects collected is at least equal to a user specified value (for example, with the sweep net method, if $\hat{\lambda}_f \geq 0.06$ TPB/sweep, then the sample is infested).

When a consistent pattern of infestation rate and risk occurs from a collection of samples that conform to both ecological and field experience, the sampling of a habitat can stop and the management decision can be made. The consultant is now hunting (S. Fleischer, *personal communication*) for insects within habitat classes rather than randomly or haphazardly sampling cotton fields.

Acquisition of field data

The scouting method for TPB adults used a standard 380 mm sweep net to estimate TPB abundance as numbers of adult bugs per sweep ($\hat{\lambda}_f$). Samples located in different habitats were selected 15 June 1999 with the support of imagery and knowledge of planting date.

A GPS reading using a Topcon Turbo-G1 (Paramus, NJ, USA) receiver was collected at each location after the last sweep. These GPS readings were post-processed (using the continuously operating reference station (or CORS) station at Memphis, TN, USA (mem2) as the base station) to remove the bias introduced by selective availability and other sources of error (Kennedy, 1996). The sites (and attributes) were placed onto the geo-referenced multi-spectral image composite using the GIS package, ArcView® 3.2a (ESRI, Redlands, CA, USA).

Results were first summarized by assuming the samples were drawn from a common population (*i.e.*, no stratification). A second summarization re-grouped the sample estimates according to (1) different cotton habitats and (2) management and weather conditions.

Results

Modeling results

Negative binomial distribution. The CDF graphs (Figures 1–3) show that the probability of observing zero insects decreases as sample unit size (s) and λ increase. For example, comparing the CDF using the smallest sized sample unit ($s = 1$) for the two lower infestation levels ($\lambda_c = 0.04$ or 0.10), the expected proportion of sample units found to be uninfested ($P(r = 0)$) is close to $1 - \lambda_c$. However, for $\lambda_c = 0.40$, the expected proportion of sample units uninfested is larger than $1 - \lambda_c$ and is slightly more than 0.67. Despite the forcing toward random dispersion, if the sample unit size remains small ($s = 1$), some individual sample units are likely to have more than one insect as the infestation rate increases (Pielou, 1977). This is in agreement with conclusions found by others (Wilson and Room, 1983; Wilson *et al.*, 1989; Willers *et al.*, 1990; Davis, 1994) and conforms to field observations (McKibben and Willers, *personal communication*).

The model also provides for the establishment of lower and upper limits for the number of insects observable in sample units of different sizes while sampling a pest population occurring at a particular infestation rate. Under the assumption of a random dispersion pattern, the model allows users to determine the smallest number of insects (r_1) observable for sample units of different sizes and infestation rates (λ_c) such that the probability of observing that value is very small ($P(r_1) \approx 0$). For example, in Figure 2 where the infestation rate is 0.1, the probability of observing 0 insects in a sample is about 0.1 if the sample size is 25 plants, but it is almost zero when the sample unit size is 50 adjacent plants. Therefore, if a sample size of 50 adjacent plants were taken and no insects were found, you could conclude for that sample the infestation rate is smaller than 0.1. However, referring back to Figure 1, the probability of observing 0 insects in a sample unit of this size is 0.13 if the infestation rate is 0.04. By comparing values for the lower limit of expected numbers of insects observable in samples with a particular sample unit size, a sampler can establish an abundance interval (*e.g.*, $0.04 \leq \lambda \leq 0.1$ when $s = 50$ and 0 insects were found) for a habitat with results from a single sample. Collecting several more samples and comparing these results to the initial range establishes the plausibility that the interval remains reasonable or needs to be adjusted.

Similarly, the model allows samplers to determine where the CDF plateaus at an upper limit (r_2) for the number of insects most likely to be collected for a choice of sample unit size and different infestation rates. For example, with a sample unit whose size is 50 adjacent plants, it is essentially impossible to capture more than 5 insects in a sample if $\lambda_c = 0.04$ (Figure 1). If instead 8 insects are captured with a sample unit size of 50 adjacent plants, we can immediately conclude that the infestation rate at the sample location is > 0.04.

Values outside an interval ($r_1 \leq r \leq r_2$) for a specified sample unit size and sampling method are unusual observations at an assumed infestation rate and indicate that either the assumed rate is incorrect or the sampled habitat is heterogeneous. Most often, while sampling cotton habitats for insects, this means that more samples should be collected. The additional samples should provide evidence to decide whether a higher (or lower) rate is more plausible or that more than one pest population occurs in a habitat that was not properly classed.

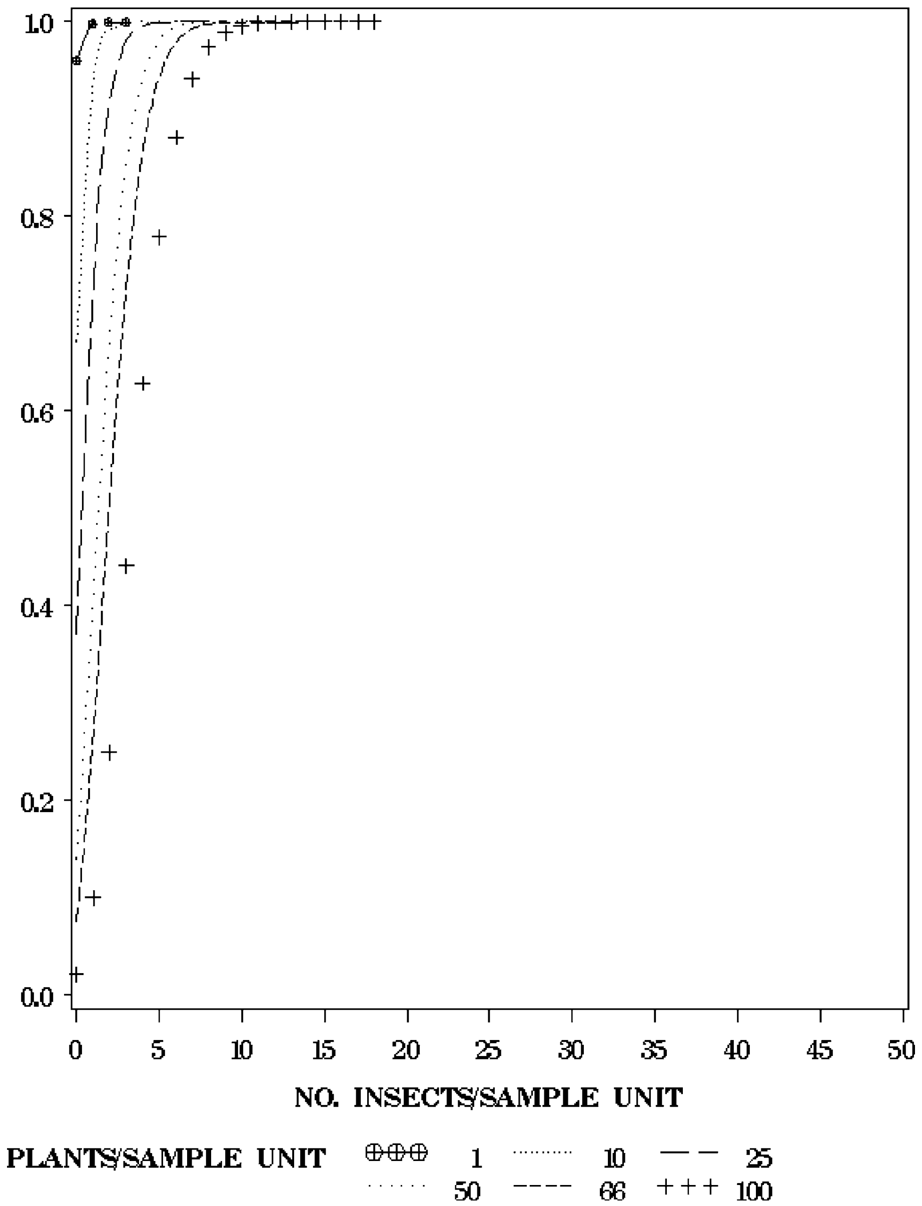

Figure 1. Cumulative distribution functions for the number of insects per sample unit as modeled by the negative binomial distribution for six different sample unit sizes (adjacent plants) where $\lambda_c = 0.04$ and a random dispersion pattern ($k = 50$).

Resampling results and mixed population densities. While improvements in detecting insects are possible by choosing sufficiently large sample unit sizes, the accuracy of estimating insect abundance is not improved by increasing the number of samples if

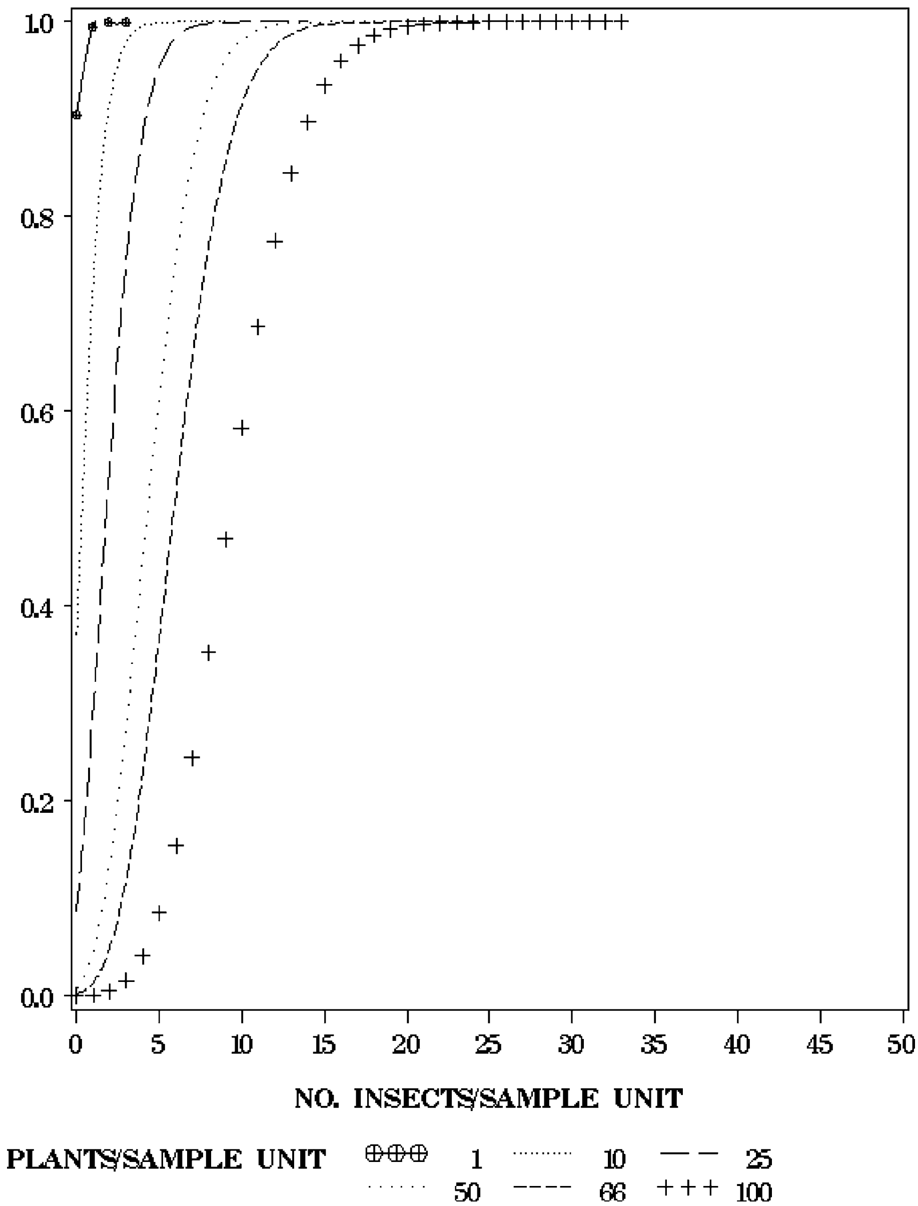

Figure 2. Cumulative distribution functions for the number of insects per sample unit as modeled by the negative binomial distribution for six different sample unit sizes (adjacent plants) where $\lambda_c = 0.10$ and a random dispersion pattern ($k = 50$).

multiple population densities exist. This simulation experiment demonstrates the poor performance of a SRS design if field samples are not stratified according to habitats. The enumeration of sample combinations provided all possible estimates of

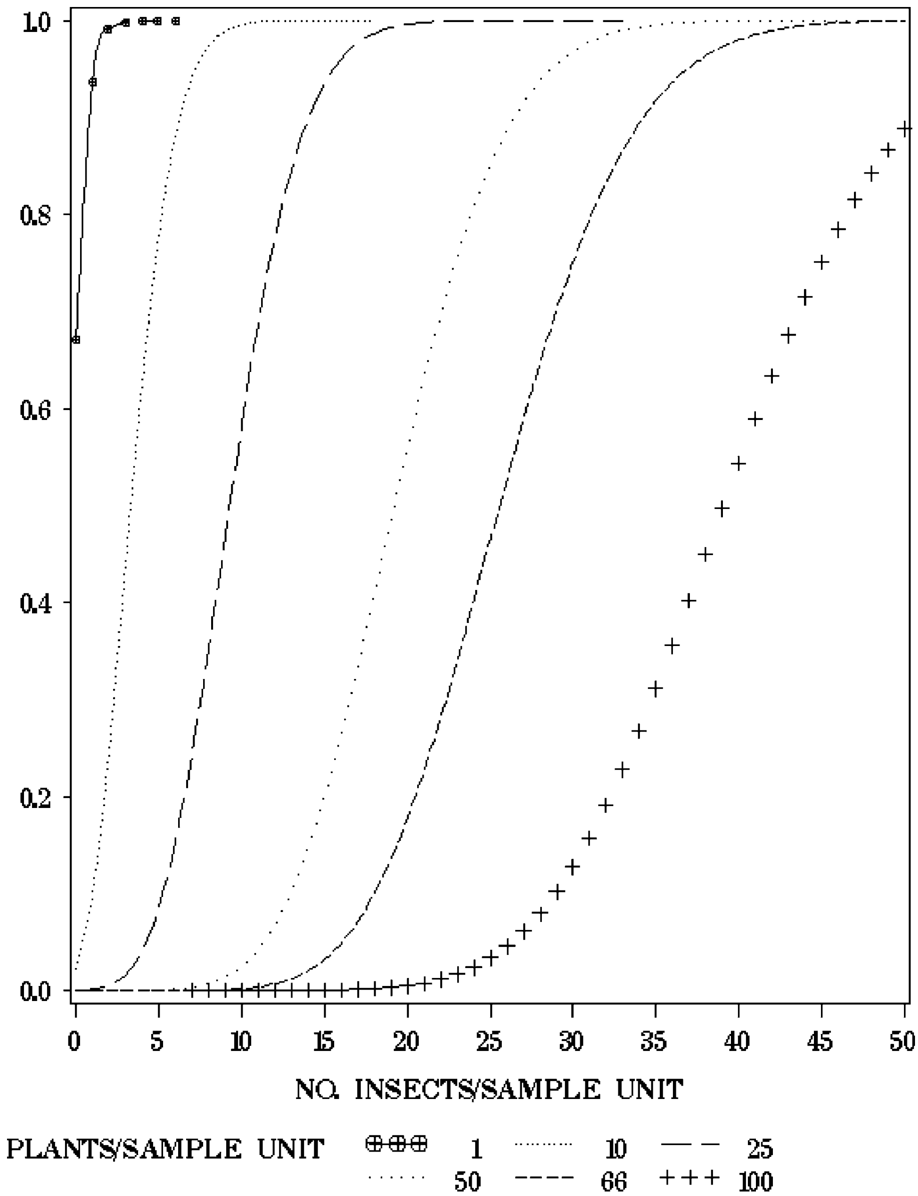

Figure 3. Cumulative distribution functions for the number of insects per sample unit as modeled by the negative binomial distribution for six different sample unit sizes (adjacent plants) where $\lambda_c = 0.40$ and a random dispersion pattern ($k = 50$).

mean pest abundance (Ash, 1993; Gonick and Smith, 1993) from the computer-generated field. Even though two non-overlapping pest densities infested the sample units in the simulated field, the resampling process modeled the incorrect assumption of a common pest density distribution.

The resampling distribution (Figure 4) indicates the modal value estimating the average infestation rate of the field is near 0.20, as would be expected by averaging the two densities of the initial set of 24 field units (since to the modeler, these values are known). The distribution of sample estimates indicates that increasing sample size will not improve estimate accuracy when the assumption of sampling a homogenous population (Thompson, 1992) is not applicable. Estimates shift higher or lower if a majority of samples are collected from either population density. Since two different population densities are actually present (but unknown to the sampler by assumption) any management decision made for the entire field would not be optimal. Decisions based upon the average infestation rate would be inappropriate for the other half of the field where pest abundance differs.

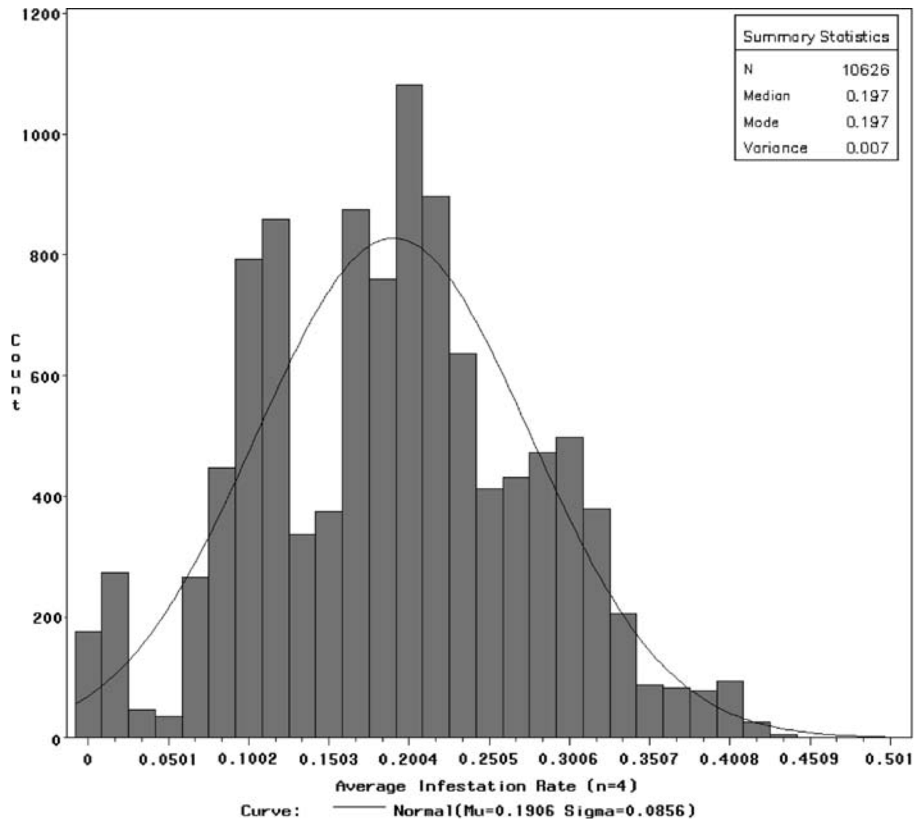

Figure 4. Frequency histogram distribution and summary statistics representing estimated mean rates of an insect pest for all combinations of 4 samples drawn from 24 choices (Table 1).

Field case study

Frequently, in commercial cotton systems neither the pest population density nor the cotton plants within fields are homogenous. Determining how many different pest population densities exist and their spatial extents requires associating the information from samples obtained 'on the ground' with information from classified, remote sensing imagery. This field example illustrates how to establish meaningful groupings among pest population density estimates with various habitats corresponding to cotton growth. The field boundary polygons defining the study area are labeled in Figure 5.

The TPB sample data ($\hat{\lambda}_f$, the actual counts and total sweeps per sample) are listed by field name (Table 3) without any information about relationships to habitat classes. When presented without a spatial context, these data exhibit distributional patterns and behaviors as if they are samples from a common population. However, even though they are grouped by field, the set of samples are too noisy and sparse to be useful. A larger sample size in each field would not improve the ability to make a management decision.

The scouting site locations were overlaid upon the multi-band image composite (Figure 6) along with auxiliary variables describing the number of TPB adults

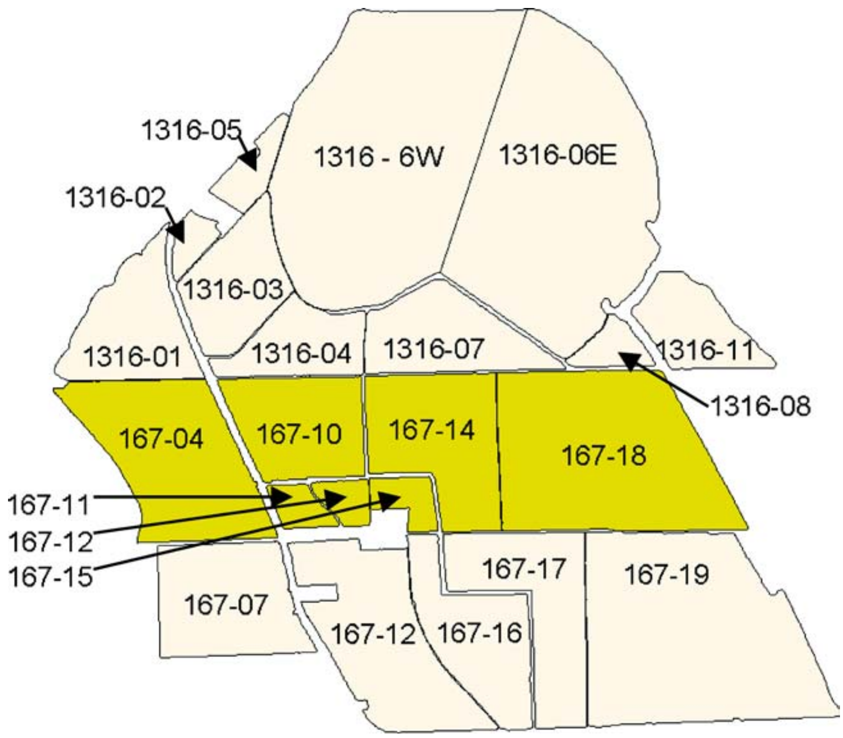

Figure 5. Field polygons of the study area and associated farm labels. The first value is the farm tract number while the second value is the field number within the tract. The bright yellow fields are where an aircraft applied ULV malathion and rainfall did not occur after the application (see text).

Table 3. Estimates ($\hat{\lambda}_f$ (in bold)) of tarnished plant bug adults per sweep on 15 June 1999 at sample locations within named field polygons (Figure 5)

167-04	1316-01	167-10	1316-04	1316-03	1316-02	1316-05	1316-07	167-14	167-18	1316-6E	1316-6W	1316-08	167-19
0.0, 0/50	**0.06**, 2/33	**0.0**, 0/50	**0.0**, 0/33	**0.10**, 5/50	**0.07**, 3/42	**0.08**, 4/50	**0.02**, 1/50	**0.0**, 0/33	**0.0**, 0/50	**0.03**, 1/33	**0.03**, 1/33	**0.12**, 4/33	**0.0**, 0/33
0.0, 0/50	**0.09**, 3/33	**0.0**, 0/50					**0.04**, 2/50		**0.0**, 0/50	**0.03**, 1/33			**0.0**, 0/33
0.0, 0/50										**0.03**, 1/33			
										0.03, 1/33			

The ratios following the estimates are the number of bugs (numerator) collected for the total number of sweeps (denominator).

collected and a qualitative assessment of beneficial abundance. For the beneficial counts, (B−) indicates none were found, (B ±) indicates only a few were found and (B +) indicates more than 6 were collected for the sample.

The multi-band image composite produced a habitat map (Figure 7) that provided information on where to sample for insects. The NDVI values of geo-referenced pixels described the variability in crop community structure across the landscape. As with insect population densities, these NDVI values [or other vegetation indices (Elvidge and Chen, 1995)] corresponded to habitats captured by 'populations of pixels' that differed in statistics for the mean, variance, kurtosis and skewness. An eight-color list of habitats was established. For the categorical cotton habitat classes (poor, marginal, good and best), plants increased in height, density, growth rate, squaring rate and canopy development. Plant *et al*. (2001) describes the NDVI with respect to cotton production. Willers *et al*. (1999) discuss the NDVI ratio with respect to habitat classification, line-intercept sampling and TPB abundance.

Knowledge of management practices at the time established that while effects on pest density were not different, the causes for plants classed as poor habitats differed among fields. For example, the red regions in Fields T1316-04, -6E, -6W and -07 did not represent cotton in the same state of development as red regions within late-planted fields. Cotton in these four northern fields was older, but grew in

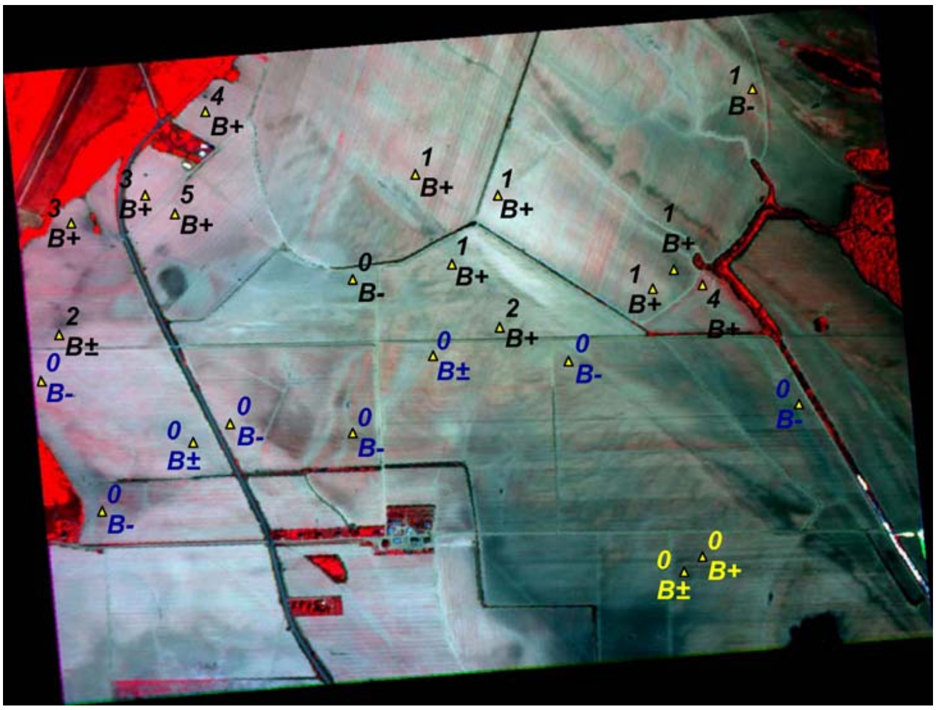

Figure 6. Tarnished plant bug counts and beneficial insect ratings from sample sites overlaid on a false color multi-band composite acquired on 8 June 1999, at Perthshire Farms, Bolivar County, MS, USA (see text).

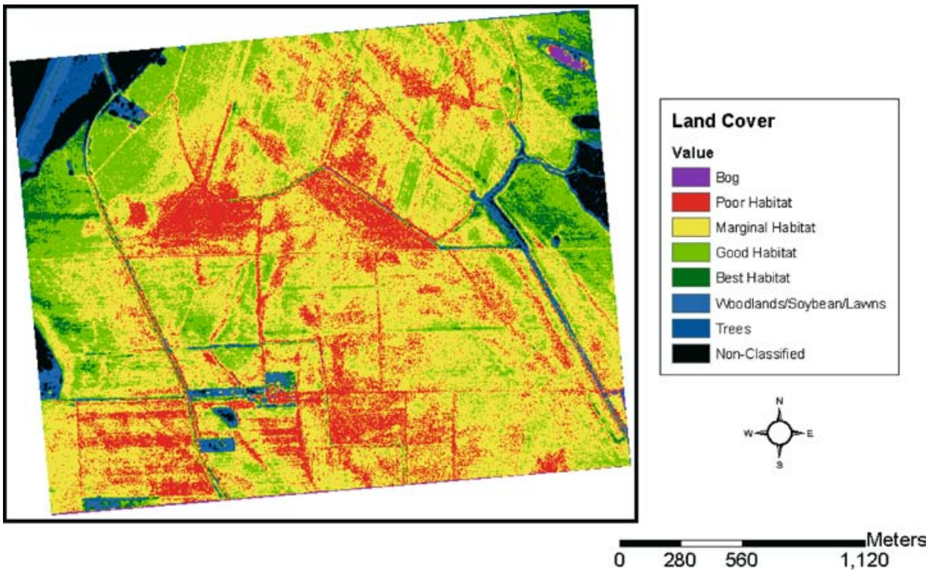

Figure 7. June 8, 1999 image, where each pixel has been assigned a normalized difference vegetation index (NDVI) value and classed into different land cover categories. Habitat classes refer to stages of development in the cotton crop.

depressional areas comprised of heavy, silty clay soils. Late-planted cotton plants in five southern fields (T167-07, -12, -16, -17 and -19) were not growing in this type of soil but were generally developing in sandy loam or sandy silt soils. While the two different regions portrayed a similar NDVI range, the causes of the similarity differed. Image classification procedures alone cannot correctly define habitat structure if local knowledge about crop and soil conditions is ignored.

The TPB sample data revealed correspondence between numbers found in sweeps and the habitat class color. The two samples in field T167-19 and the sample in T1316-04 were zero valued as expected. In the first case, the crop was not mature enough to have squares favorable for attraction of TPB. In the second case, even though the crop was planted earlier, the cotton sampled in this area was delayed due to water stress and had no squares. TPB prefer cotton plants with squares; hence, from ecological considerations only a few samples are necessary where the habitat is poor (of any age) or marginal (and late-planted). However, for other habitats comprised of older cotton, spraying and rainfall information were necessary to best allocate sample sites to management decision categories. For example, the sequence of 8, zero-valued, samples for fields T167-04, -10, -14 and -18 did not match expectations based on examination of the habitat map alone. These samples were from habitats favorable (*e.g.*, green and dark green) for establishment of TPB. This sequence of zeros was not due to sampling or observer error, because enough sweeps were used to create large enough sized sample units (Tables 2 and 3) such that if $\lambda_f \geq 0.06$, confidence would be high that at least 1 insect would be found. These central fields (Figure 5, shaded in yellow), where many non-infested samples were found in favorable habitats, had older plants treated with ULV malathion by air and no rain fell soon after the application (Kenneth Hood, *personal communication*). In contrast, in the upper region of Figure 6, where the majority of infested samples resided, rain fell several hours after an ULV malathion application by air. Malathion is highly soluble in water (Nemec and Adkisson, 1969; El-Lissy *et al*., 1997). This application was rinsed off by rain before effectively controlling resident TPB populations (or eliminating beneficials) in the habitats. For these samples, the TPB population densities were more severe in the good and best habitats, occurred in low numbers in the marginal habitat and were absent from the poor habitat.

The information just discussed can be summarized as a decision table (Table 4). Only those fields in the northern area of the image needed to be re-treated for TPB and only where the good and best habitats occurred ($\hat{p} > 0.7$). Some marginal areas of the northern fields did have sample estimates of TPB close to the action threshold, but risk was low ($\hat{p} \approx 0$). The decision was made to not re-treat this class, but wait and see what happens. The poor habitat areas were left alone. On the other hand, the application of the middle fields (where it did not rain following the spray) was effective ($\hat{p} \approx 0$) and none of these habitats needed to be retreated. Plant bugs were not a problem ($\hat{p} \approx 0$) for the southern most fields not yet sprayed for the first time, because the cotton plants were still young and not attractive for colonization. Samples from the various habitats and historical knowledge spatially answered the general question, "Is there an insect problem here?" in different ways because estimates for risk differed among them.

A spatial assessment of TPB abundance with relatively small sample sizes in each habitat class was accomplished for an area of 218 ha. For other years, as weather influences vary or mean pest abundance in the landscape becomes very high, different management categories need to be made. At times, the samples may indicate that all habitats are at risk. If such is the case, the methods and principles described still apply.

Table 4. The sweep net sample estimates[a] from Table 3 for tarnished plant bugs (TPB) grouped into habitat, crop planting date and spray event categories

	First planting		Later planting	
Category/habitat class	Sprayed (No rain)	Sprayed (Rain)	Sprayed	Not sprayed
Best	NS	**0.12** ($\hat{p} = 1.0$)	NA	NA
Good	**0.0, 0.0, 0.0, 0.0, 0.0** ($\hat{p} = 0.0$)	**0.03, 0.06, 0.07, 0.08, 0.09, 0.10** ($\hat{p} = 0.83$)	NA	NA
Marginal	**0.0, 0.0, 0.0** ($\hat{p} = 0.0$)	**0.02, 0.03, 0.03, 0.03, 0.03, 0.04** ($\hat{p} = 0.0$)	NA	**0.0, 0.0** ($\hat{p} = 0.0$)
Poor	NS	**0.0** ($\hat{p} = 0.0$)	NA	NS

The risk estimates ($\hat{p}$) indicate that two habitat-management categories need to be promptly retreated since $\hat{p} > 0.7$. One 'marginal' habitat-management category will not need to be retreated, but should be monitored as the TPB population located there is near the action threshold.

NS ≡ Not Sampled; NA ≡ Not Applicable; Sprayed ≡ ULV Malathion (highly water-soluble).

[a]The action threshold (using a sweep net) at this time of the cotton production season is $\hat{\lambda}_f \geq 0.06$ TPB/sweep.

Discussion

Field experience and model results (Figures 1–3) show (Table 2) that having a sample unit size of no more 0.91 m for more than four adjoining rows (about 9 contiguous plants per row) is sufficient when sampling by drop cloth (Willers and Akins, 2000; Willers *et al.*, 2000). For visual quadrat samples (Willers *et al.*, 1990; Williams *et al.*, 1996), the sample unit should be comprised of about 50 adjacent plants (or 2 rows by 2.74 m) but to save sampling time, most often 5–10 of these plants (selected according to the sampler's judgment) are used to estimate pest abundance of the sample. For sweep net samples, the sample unit most often used is 33 sweeps or a sample unit size near 66 plants (assuming 2 plants are bisected by a sweep). When a sweep net is used, the sweeps are collected down a cotton row until nearly one-half of a sample is collected. One then moves to an adjoining row and returns back toward the beginning point to complete the sample.

There are several benefits to using probability models as opposed to field studies for developing and evaluating a sampling design. It is laborious and expensive to compare various sampling designs in the field. Field results are highly biased by intrinsic edaphic factors (Daubenmire, 1974) and differences in observer skill (Harrington, 1987; Schaalje and Butts, 1992). Observers also cause disturbances to the canopy and insect fauna, so it is impossible to compare two sampling designs at exactly the same site. Biases in results also occur because the parametric density and dispersion pattern of insects in a field are not known. In contrast, by first making judgments about the performance of a sampling design using models (Willers *et al.*, 1990) or resampling methods (Hutchison *et al.*, 1988; Naranjo and Hutchison, 1997; Willers *et al.*, 2000), it is possible to improve the interpretation of sample estimates from the field if the sample sizes are small and the sample unit sizes are larger than single plants.

The resampling experiment demonstrated that larger sample unit sizes with large sample sizes did not improve the accuracy of an SRS estimator when the sample universe was comprised of mixed (D'Agostino and Stephens, 1986) pest population density distributions. There is a tradition among entomologists (and others), which holds that accuracy of an estimate improves if the sample size is increased (Karandinos, 1976; Fleischer *et al.*, 1999). The probability of observing zero insects for different sample unit sizes (Figures 1–3) and cases where mixed population densities exist (Figure 4; Table 4) should make field practitioners of integrated pest management (IPM) pause for a moment of reflection. Even with larger sample unit sizes, accuracy of an estimate of pest abundance (Ludwig and Reynolds, 1988; Buntin, 1994) will not improve with greater numbers of samples if multiple population densities exist in the sampling universe. The ability to estimate pest abundance only becomes worse if sample unit sizes are too small.

To make better management decisions, the sampling effort must be stratified, but efficiently defining relevant strata on large commercial farms is difficult without remote sensing. Multi-spectral, remote sensing imagery of high spatial resolution (Jensen, 1986; Pouncey *et al.*, 1999) precisely defines the number and inter-relationships among cotton habitats (or sampling strata). Pest abundance often corresponds to differences in cotton habitat structure (Willers *et al.*, 1999, 2000).

Many sampling plans strive to save time (Jones, 1994; Gutierrez, 1996). Large commercial farms place a premium upon saving time. The use of remote sensing imagery saves sample time because an image-based, habitat stratified, SRS design provides a practical correspondence between the sizes of the sample units and image pixels. The pixel size defines the ground sample distance (GSD) covered by the field of view of a single detector within a sensor array (Anonymous, 1997). The GSD defines the number of pixels needed to capture a cotton field in a digital image. These pixels in a high spatial resolution image ($1\ m^2 \leq$ pixel $\leq 4\ m^2$) conveniently define the size and location of the minimal 'habitable space' (Pedigo, 1994) of the insect pest. Since the pixel is scalable to the sample unit size, by analogy, the pixels define the sampling universe (Thompson, 1992; Pedigo, 1994) of the cotton field. The geographic label attaches samples within a habitat class to specific pixels in a geo-referenced image. Associating the management decision (and other auxiliary variables) of the samples to the pixels for the rest of the habitat class creates the mechanism to build spatial pesticide prescriptions (Dupont *et al*., 2000; Reed, 2001a b c; Seal *et al*., 2001) using GIS techniques.

Conclusions

Most sampling techniques developed prior to the 1990's cannot provide estimates of cotton insect pest abundance at fine spatial scales over a large landscape. For large commercial production fields scattered over a large farm landscape, there is not enough available time or labor at an affordable cost to acquire enough samples to build a reliable map of insect abundance. Larger sized sample units aid in minimizing the observance of zeros in samples when the pest density is at or above an action threshold. Optimal sampling in the presence of heterogeneous pest population densities with small sample sizes requires habitat delineation using imagery. The capability to stratify samples by cotton habitats improves the ability to segregate statistical distributions representing different population densities of fruit-feeding pest insects and reduces the bias and variance of sample estimates. Knowledge of past management practices and weather also assist experienced consultants in establishing spatial boundaries across habitat classes where a pest is (1) of no concern, (2) potentially a problem or (3) severe enough to take action. The integrated sampling system estimated cotton fruit feeding insect population densities across many fields with surprisingly small sample sizes.

Acknowledgments

We thank Drs. Blake Layton, Randy Luttrell, L. Wes Burger, Mr. Charles Hill and several anonymous reviewers for comments and suggestions on the initial drafts. The Advanced Spatial Technologies for Agriculture Program and the GeoResources Institute, Mississippi State University, have partially provided funds and technical support for the work described. This article reports the results of research only. Mention of a trademark, proprietary product or vendor does

not constitute guarantee or warranty of the product by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that may also be suitable.

References

Andrewartha, H. G. and Birch, L. C. 1970. The Distribution and Abundance of Animals. (University of Chicago Press, Chicago, USA) 782p.

Anonymous, 1997. Multispectral Imagery Reference Guide. (Logicon Geodynamics, Fairfax, VA, USA) 200p.

Anscombe, F. J. 1949. The statistical analysis of insect counts based on the negative binomial distribution. Biometrics **5**, 165–173.

Anscombe, F. J. 1950. Sampling theory of the negative binomial and logarithmic series distributions. Biometrika **37**, 358–382.

Ash, C. 1993. The Probability Tutoring Book. An Intuitive Course for Engineers and Scientists (and Everyone Else). (IEEE Press, Piscataway, NJ, USA) 470p.

Beyer, W. H. 1968. Handbook of Tables for Probability and Statistics, 2nd edn. (CRC Press, Cleveland, OH, USA) 642p.

Buntin, G. D. 1994. Developing a primary sampling program. In: *Handbook of Sampling Methods for Arthropods in Agriculture*, edited by L. P. Pedigo and G. D. Buntin (CRC Press, Boca Raton, FL, USA), pp. 99–115.

Byerly, K. F., Gutierrez, A. P., Jones, R. E. and Luck, R. F. 1978. A comparison of sampling methods for some arthropod populations in cotton. Hilgardia **46**, 257–282.

Colinvaux, P. A. 1973. Introduction to Ecology. (John Wiley & Sons, New York, USA) 621p.

D'Agostino, R. B. and Stephens, M. A. 1986. Goodness-of-Fit Techniques. (Marcel Dekker, New York, USA) 560p.

Daubenmire, R. F. 1974. Plants and Environment: A Textbook of Autecology, 3rd edn. (John Wiley & Sons, New York, USA) 442p.

Davis, P. M. 1994. Statistics for describing populations. In: *Handbook of Sampling Methods for Arthropods in Agriculture*, edited by L. P. Pedigo and G. D. Buntin (CRC Press, Boca Raton, FL, USA), pp. 33–54.

Dupont, J. K., Campanella, R., Seal, M. R., Willers, J. L. and Hood, K. B. 2000. Spatially variable insecticide applications through remote sensing. In: *2000 Proceedings of the Beltwide Cotton Conferences*, edited by P. Dugger and D. Richter **2** (National Cotton Council, Memphis, TN, USA), pp. 426–429.

Edirisinghe, A., Chapman, G. E. and Louis, J. P. 2001. Radiometric corrections for multispectral airborne video imagery. Photogrammetric Engineering and Remote Sensing **67**(8), 915–922.

El-Lissy, O., Shepherd, W. and Meyers, F. 1997. Longevity of malathion ULV applications against boll weevil under different weather conditions in the coastal bend of Texas. In: *1997 Proceedings of the Beltwide Cotton Conferences*, edited by P. Dugger and D. Richter **2** (National Cotton Council, Memphis, TN, USA), pp. 1209–1211.

Elvidge, C. D. and Chen, Z. 1995. Comparison of broad-band and narrow-band red and near-infared vegetation indices. Remote Sensing and Environment **54**, 38–48.

ESRI Institute. 1994. Map Projections. Georeferencing Spatial Data, (Environmental Systems Research Institute, Redlands, CA, USA).

Fleischer, S. J., Blom, P. E. and Weisz, R. 1999. Sampling in Precision IPM: When the objective is a map. Phytopathology **89**(11), 1112–1118.

Gonick, L. and Smith, W. 1993. The Cartoon Guide to Statistics. (Harper Collins, New York, USA) 230p.

Gutierrez, A. P. 1996. Applied Population Ecology. A Supply-Demand Approach. (John Wiley and Sons, New York, USA) 300p.

Harrington, R. 1987. Varying efficiency in a group of people sampling cabbage plants for aphids (Hemiptera: Aphididae). Bulletin Entomological Research **77**, 497–501.

Hutchison, W. D., Hogg, D. B., Poswall, M. A., Berberet, R. C. and Cuperus, G. W. 1988. Implications of the stochastic nature of Kuno's and Green's fixed-precision stop-lines: Sampling plans for the pea aphid (Homoptera: Aphididae) in alfalfa as an example. Journal Economic Entomology **81**, 749–758.

Jensen, J. R. 1986. Introductory Digital Image Processing, A Remote Sensing Perspective. (Prentice-Hall, Englewood Cliffs, NJ, USA) 379p.

Jones, V. P. 1994. Sequential estimation and classification procedures for binomial counts. In: *Handbook of Sampling Methods for Arthropods in Agriculture*, edited by L. P. Pedigo and G. D. Buntin (CRC Press, Boca Raton, FL, USA), pp. 175–205.

Karandinos, M. G. 1976. Optimum sample size and comments on some published formulae. Bulletin Entomological Society America **22**, 417–421.

Kennedy, M. 1996. The Global Positioning System and GIS: An Introduction. (Ann Arbor Press, Chelsea, MI, USA) 268p.

Krebs, C. J. 1978. Ecology: The Experimental Analysis of Distribution and Abundance. 2 edn. (Harper and Row, New York, USA) 678p.

Ludwig, J. A. and Reynolds, J. F. 1988. Statistical Ecology: A Primer on Methods and Computing. (John Wiley & Sons, New York, USA) 337p.

Manly, B. F. J. 1997. Randomization, Bootstrap, and Monte Carlo Methods in Biology, 2 edn. (Chapman and Hall, London, UK) 399p.

Morris, R. F. 1955. The development of sampling techniques for forest insect defoliators, with particular reference to the spruce budworm. Canadian Journal Zoology **33**, 225–294.

Naranjo, S. E. and Hutchison, W. D. 1997. Validation of arthropod sampling plans using a resampling approach: Software and analysis. American Entomologist **43**(1), 48–57.

Nemec, S. J. and Adkisson, P. L. 1969. Effects of simulated rain and dew on the toxicity of certain ultra volume insecticidal formulations. Journal Economic Entomology **62**, 71–73.

Pedigo, L. P. 1994. Introduction to sampling arthropod populations. In: *Handbook of Sampling Methods for Arthropods in Agriculture*, edited by L. P. Pedigo and G. D. Buntin (CRC Press, Boca Raton, FL, USA), pp. 1–11.

Pedigo, L. P. and Buntin, G. D. (eds). 1994. Handbook of Sampling Methods for Arthropods in Agriculture. CRC Press, Boca Raton, FL, USA, 714p.

Pielou, E. C. 1977. Mathematical Ecology. (John Wiley and Sons, New York, USA) 384p.

Pinter, P. J. Jr., Hatfield, J. L., Schepers, J. S., Barnes, E. M., Moran, M. S., Daughtry, C. S. T. and Upchurch, D. R. 2003. Remote sensing for crop management. Photogrammetric Engineering and Remote Sensing **69**, 647–664.

Plant, R. E., Munk, D. S., Roberts, B. R., Vargas, R. N., Travis, R. L., Rains, D. W. and Hutmacher, R. B. 2001. Application of remote sensing to strategic questions in cotton management and research. Journal Cotton Science **5**, 30–41.

Pouncey, R., Swanson, K. and Hart, K. (ed.) 1999. ERDAS Field Guide, 5th ed. (ERDAS, Atlanta, USA), 672p.

Pritsker, A. B. and Pegden, C. D. 1979. Introduction to Simulation and SLAM. (Halsted Press, John Wiley and Sons, New York, USA) 588p.

Reed, J. 2001a. Dollars in and dollars out, Part 1. Cotton Farming **45**(7), 26–39.

Reed, J. 2001b. Dollars in and dollars out, Part 2. Cotton Farming **45**(8), 27–29.

Reed, J. 2001c. Dollars in and dollars out, Part 3. Cotton Farming **45**(9), 6–8.

Richards, J. A. and Jia, X. 1999. Remote Sensing Digital Image Analysis. An Introduction. 3rd edn. (Springer-Verlag, Berlin, Germany) 363p.

Rouse, J. W. Jr., Haas, R. H., Deering, D. W., Schell, J. A. and Harlan, J. C. 1974. Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation. NASA/GSFC Type III Final Rpt. (Greenbelt, MD, USA).

Schaalje, G. B. and Butts, R. A. 1992. Binomial sampling for predicting density of Russian Wheat Aphid (Homoptera: Aphididae) on winter wheat in the fall using a measurement error model. Journal Economic Entomology **85**, 1167–1175.

Seal, M., Dupont, K., Bethel, M., Lewis, D., Johnson, J., Willers, J., Hood, K., Hardwick, J., Leonard, R. and Bagwell, R. 2001. Utilization of remote sensing technologies in the development and implementation of large-scale spatially-variable insecticide experiments in cotton. In: *2001 Proceedings of the*

Beltwide Cotton Conferences, edited by P. Dugger and D. Richter **2** (National Cotton Council, Memphis, TN, USA), pp. 1010–1018.

Southwood, T. R. E. 1978. Ecological Methods with Particular Reference to the Study of Insect Populations. (Halsted Press, John Wiley & Sons, New York, USA) 524p.

Thenkabail, P. S., Smith, R. B. and DePauw, E. 2002. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. Photogrammetric Engineering and Remote Sensing **68**(6), 607–621.

Thompson, S. K. 1992. Sampling. (Wiley-Interscience, New York, USA) 343p.

Willers, J. L. and Akins, D. S. 2000. Sampling for tarnished plant bugs in cotton. Southwestern Entomologist. Supplement No **23**, 39–57.

Willers, J. L., Boykin, D. L., Hardin, J. M., Wagner, T. L., Olson, R. L. and Williams, M. R. 1990. A simulation study on the relationship between the abundance and spatial distribution of insects and selected sampling schemes. In: *Proceedings of the 1990 Kansas State University Conference on Applied Statistics in Agriculture*, edited by G. A. Milliken and J. R. Schwenke (Kansas State, Manhattan, KS, USA), pp. 33–45.

Willers, J. L., Ladner, W. L., McKinion, J. M. and Cooke, W. H. 2000. Application of computer intensive methods to evaluate the performance of a sampling design for use in cotton insect pest management. In: *Proceedings of the 2000 Kansas State University Conference on Applied Statistics in Agriculture*, edited by G. A. Milliken (Kansas State, Manhattan, KS, USA), pp. 119–133.

Willers, J. L., Seal, M. R. and Luttrell, R. G. 1999. Remote sensing, line-intercept sampling for tarnished plant bugs (Heteroptera: Miridae) in Mid-south cotton. Journal Cotton Science **3**, 160–170.

Williams, M. R. 1998. Cotton Insect Losses-1997. In: *1998 Proceedings of the Beltwide Cotton Conferences*, edited by P. Dugger and D. Richter **2** (National Cotton Council, Memphis, TN, USA), pp. 904–925.

Williams, M. R., Wagner, T. L. and Willers, J. L. 1996. Revised Protocol for Scouting Arthropod Pests of Cotton in the Midsouth. Mississippi Agriculture Forestry Experiment Station Technical Bulletin 206. (Mississippi State University, Mississippi State, MS, USA) 28p.

Wilson, L. T. and Room, P. M. 1983. Clumping patterns of fruit and arthropods in cotton, with implications for binomial sampling. Environmental Entomology **12**, 50–54.

Wilson, L. T., Sterling, W. L., Rummel, D. R. and DeVay, J. E. 1989. Quantitative sampling principles in cotton IPM. In: *Integrated Pest Management Systems and Cotton Production*, edited by R. E. Frisbie, K. M. El-Zik and L. T. Wilson (John Wiley & Sons, New York, USA), pp. 85–119.

Yuan, D. and Elvidge, C. D. 1996. Comparison of relative radiometric normalization techniques. ISPRS Journal Photogrammetric Remote Sensing **51**, 117–126.